

**SPECIFICATION <EXCERPT>**

[0054]

The number of unrolling loops that is obtained from the function FUNC described above is also registered in the loop control information table as loop control information. The loop transformation unit 30 receives that loop control information table, and executes various kinds of optimization according to that information. For example, various kinds of optimization are executed or inhibited according to information in the table such as loop unrolling/execute/expand 5 times, loop reduction/execute, and software pipelining/inhibit.

[0055]

FIG. 4 is a drawing showing an example of optimization according to an embodiment of the present invention; FIG. 5 is a drawing explaining an example of the operation of the loop optimization unit of an embodiment of the present invention; and FIG. 7 is a drawing showing an example of a table that is used in an embodiment of the present invention.

[0056]

How the loop optimization unit 14 of an embodiment of the present invention operates on the input of an actual program is explained below using program A shown in FIG. 4(A) as an example.

[0057]

When program A is given to the compiler, program A is inputted to the loop optimization after passing through the front end unit 12 and the program structure analysis unit 13, and finally reaches the loop optimization control unit 20.

[0058]

As shown in FIG. 5, in the loop optimization control unit 20, first,

the option information analysis unit 21 operates, and recognizes according to an option specification that a high optimization level is required, as well as recognizes that the renaming inhibit option has been specified. In addition, it also recognizes linking optimization at the high optimization level. In order to transmit this information to a later phase, the option information analysis unit 21 creates an option information table as shown in FIG. 6(A)

[0059]

The optimization levels in the option information table are divided as follows:

0: No optimization

1: Low-level optimization

2: Medium-level optimization

3: High-level optimization

Moreover, optimization in the option information table is categorized as follows:

[0060]

0x0001: Loop unrolling

0x0002: Loop reduction

0x0004: Software pipelining

0x0008: Renaming

In other words, 0x0007 indicates that loop unrolling, loop reduction and software pipelining are the object of processing.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-101776

(43) 公開日 平成8年(1996)4月16日

(51) Int.Cl.<sup>4</sup>

識別記号

庁内整理番号

F I

技術表示箇所

G 0 6 F 9/45

7737-5B

G 0 6 F 9/44

3 2 2 G

審査請求 未請求 請求項の数 8 O L (全 15 頁)

(21) 出願番号 特願平6-238670

(22) 出願日 平成6年(1994)9月30日

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区小田中1015番地

(72) 発明者 原口 正寿

神奈川県川崎市中原区小田中1015番地

富士通株式会社内

(72) 発明者 林 正和

神奈川県川崎市中原区小田中1015番地

富士通株式会社内

(72) 発明者 中平 直司

神奈川県川崎市中原区小田中1015番地

富士通株式会社内

(74) 代理人 弁理士 小笠原 吉義 (外 2 名)

最終頁に続く

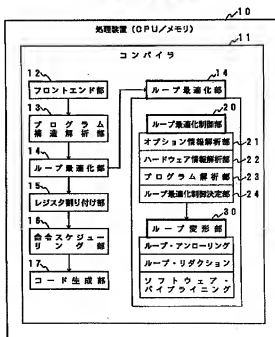
(54) 【発明の名称】 コンパイラ装置およびループ最適化制御方法

(57) 【要約】

【目的】 コンパイラ装置およびループ最適化制御方法に関し、複数のループ最適化を個々に実施するのではなく、実施する価値のあるループ最適化をより性能が上がる条件で、かつより性能の上がる組合せで実施することを目的とする。

【構成】 ループ最適化制御部20により、プログラム/ハードウェアの特性を解析し、解析結果に基づいて、同時に動作してはいけなく、あるいは同時に動作するためには制限を加える必要がある最適化群を認識することにより、複数の最適化項目についてそれぞれ実施するかしないかを決定し、また実施する場合における実施条件および最適化の実施順序を決定する。その後ループ変形部30によって、実施する最適化項目について決定された実施順序および実施条件に従ってループ最適化を行う。

本発明の原発説明図



1

## 【特許請求の範囲】

【請求項1】 ソースプログラムを解析し、最適化処理、レジスタ割り付け処理、命令スケジューリング、コード出力によってオブジェクトプログラムを出力するコンパイラ装置において、プログラム/ハードウェアの特性を解析し、同時に動作してはいけない、あるいは同時に動作するためには制限を加える必要のある最適化群を認識し、複数の最適化項目に対するループ最適化を統括的に制御するループ最適化制御部と、このループ最適化制御部の情報をもとに、個々のループ最適化を実施するループ変形部とを備えたことを特徴とするコンパイラ装置。

【請求項2】 ソースプログラムを解析し、最適化処理、レジスタ割り付け処理、命令スケジューリング、コード出力によってオブジェクトプログラムを出力するコンパイラ装置におけるループ最適化制御方法において、プログラム/ハードウェアの特性を解析する過程と、解析結果に基づいて、同時に動作してはいけない、あるいは同時に動作するためには制限を加える必要のある最適化群を認識することにより、複数の最適化項目についてそれぞれ実施するかしないかを決定し、また実施する場合における実施条件および最適化の実施順序を決定する過程と、実施する最適化項目について決定された実施順序および実施条件に従ってループ最適化を行う過程とを有することを特徴とするループ最適化制御方法。

【請求項3】 請求項2記載のループ最適化制御方法において、前記プログラム/ハードウェアの特性を解析する過程では、各演算の実行時間の重み付きでカウントされたループ内演算数、ループを変形する最適化の実施予想、ループ最適化を施した後のループ内で使用されるレジスタの見積り数をプログラム解析情報として、指定された最適化レベルに依存してプログラム解析レベルを変動させることを特徴とするループ最適化制御方法。

【請求項4】 請求項3記載のループ最適化制御方法において、ループ内で使用される変数の依存関係を解析し、その情報から、実際にループ変形を行う時と同じ条件で、各最適化が実施可能であるかどうかを判定することを特徴とするループ最適化制御方法。

【請求項5】 請求項4記載のループ最適化制御方法において、最適化予想結果を加味して、ループ変形後におけるループ内で使用されるレジスタ数を見積り、そのレジスタ数から算出したアンローリング数と、ループ内の各演算の実行時間の重みを加味した総演算数から算出したアンローリング数と、最適化の予想とに基づいてアンローリング数を決定することを特徴とするループ最適化制御方法。

【請求項6】 請求項5記載のループ最適化制御方法において、ソフトウェア・パイプラインが予想された時のアンローリング数がターゲットであるマシンの浮動

2

小数点型の除算を除く四則演算中の最大実行時間の倍数であること、かつ、前記算出された見積りレジスタ数を考慮したアンローリング数以下であることを満足するアンローリング数にすることを特徴とするループ最適化制御方法。

【請求項7】 ソースプログラムを解析し、最適化処理、レジスタ割り付け処理、命令スケジューリング、コード出力によってオブジェクトプログラムを出力するコンパイラ装置におけるループ最適化制御方法において、ループ内の他の演算に定義も参照も現れない変数が定義オペランドと参照オペランドに現れる演算のアンローリングに対して、その定義オペランドと参照オペランドの変数名を新しい変数名に置き換え、ループ外でその補正処理を施すことによりループ内の演算の並列度を高めるループ変形を行うことを特徴とするループ最適化制御方法。

【請求項8】 請求項2記載のループ最適化制御方法において、前記複数の最適化項目に、ループ・アンローリングによるループ変形と、ループ内の他の演算に定義も参照も現れない変数が定義オペランドと参照オペランドに現れる演算のアンローリングに対して、その定義オペランドと参照オペランドの変数名を新しい変数名に置き換え、ループ外でその補正処理を施すことによりループ内の演算の並列度を高めるループ変形とが含まれることを特徴とするループ最適化制御方法。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】 本発明は、最適化機能を有するコンパイラの最適化効果を促進させるコンパイラ装置およびループ最適化制御方法に関するものである。

【0002】 近年、コンピュータシステムの高速化の要求に伴い、1サイクルで複数の命令を同時に実行できるプロセッサやベクトル演算が可能なマシンが数多く開発されている。それに伴い、それらのマシンを有効に運用するために、コンパイラに対する要求も高まってきている。特に、プログラムにおけるループ部分は実行時間の比重が高いため、ループ内部のコードを最適に生成することが翻訳/実行時間の短縮に大きく貢献することになる。本発明は、ループで行われる種々の最適化を有効に活用するためのループ最適化制御方法に関するものであり、産業上の利用性は広い。

## 【0003】

【従来の技術】 従来のコンパイラ装置におけるループ最適化として、ループ・アンローリング、ソフトウェア・パイプライン等、個々のループ最適化はいくつか存在する。例えば、ループ・アンローリングに関しては、回転数をループ内で使用される浮動小数点型および整数型のレジスタ数とループ内で使用される演算数からアンローリング数を算出する方法が知られている（特開平4-307624号公報、特開平4-307625号公

3

報)。また、ループ内のそれぞれの演算に対して各演算の実行時間重みを付加してループ内演算数を算出し、その値からアンローリング数を算出する方法もすでに提案されている。

【0004】図9ないし図12は、従来のループ最適化の例であって、特に図9はループ・アンローリング、図10はリネーミング、図11および図12はソフトウェア・パイプラインの例を示す。

【0005】ループ・アンローリングは、例えば図9に示すように、ループ変数Iを含む演算を複数個に展開し、ループ回数を少なくする最適化である。リネーミングは、図10(A)に示すようなプログラム(X1~X5は異なる任意の変数)に対して、演算間の依存性をできるだけ少なくするため、変数名を変えても結果に影響のない変数を抽出して、他の変数名に置き換える最適化である。図10の例の場合、1番目と2番目の演算における変数名Aを変数名Bに変え、3番目と4番目の演算における変数名Aを変数名Cに変えることにより、①B=X1, X2=Bと②C=X3, X4=Cと③A=X5の演算間の依存性をなくしている。なお、ここでいうリネーミングとは、レジスタ割り当てあるいは命令スケジューリングで行われるレジスタ・リネーミングではなく、最適化フェーズの中でも比較的前段階で行われる中間言語に対するリネーミングである。

【0006】ソフトウェア・パイプラインによる最適化は、以下のとおりである。例えば、図11(A)に示す命令列がソースプログラムから生成されたあるループを構成する命令列であったとする。ループが図11(A)のように4回アンローリングされた場合、図11(B)に示すように4回転分の命令列を一まとまりとして処理する。これらの命令列を通常のようにスケジューリングすると、図12(A)に示すようになる。これをプロローグ部とエピローグ部に分けてパイプライン化すると図12(B)に示すようになる。さらにリスケジューリングすると、図12(C)に示すようになる。これにより、ループ内の演算にかかる時間を短縮することができ、図12(A)では、ループ内の演算に7τかかっていたものが、図12(B)では、4τに短縮されている。

【0007】以上のように、各種のループ最適化の技術が知られているが、従来、種々のループ最適化をプログラム/ハードウェアの特性を把握して統制的に制御する方法は考えられていなかった。したがって、パターンマッチ的に最適化条件を満たすというだけで、種々の最適化を動作させる傾向にあった。このため、理論的には実行性能を向上させる最適化であっても有効に活用できず、かえって性能が落ちることさえあった。ループ・アンローリングのループ展開しすぎによるスピル生成などは最適化の逆効果としてよく知られる例である。

【0008】なお、スピルとはレジスタ不足によるメモ

4

リへのロード、ストアのことをいう。プログラムで使用される変数、およびコンパイラの内部で生成される変数には、高速化のためできるだけレジスタが割り当てられる。しかし、レジスタの数に制限があり、レジスタが不足したとき、コンパイラではメモリ上の領域を利用する。メモリへのアクセスはロード、ストア命令を介して行われるため、レジスタを利用する場合と比較して大変遅く、性能低下を及ぼす原因となる。

【0009】また、従来の技術では、ループの並列化を促す最適化がループ・アンローリング、ソフトウェア・パイプライン、リネーミング程度しかなかった。特に、近年スーパースカラやVLIWというようなアーキテクチャが広く用いられるようになってきており、このようなアーキテクチャに適した新しい最適化の技術が望まれる。

【0010】

【発明が解決しようとする課題】従来の技術では以下のような問題があった。

1. オプション解析情報、ハードウェア解析情報、プログラム解析情報を統制的に管理し、その情報をもとに個々の最適化を制御する部分がないため、個々の最適化がそれぞれの実施条件を満足するというだけで実施されてしまう。このため、かえって性能が低下する可能性がある。

【0011】2. 上記1に関連して、それぞれは意味のある最適化であるが、それらが同時に動作することによってかえって性能が落ちる場合がある。また、優先度をつけて制御(抑止)すべき最適化もある。

【0012】3. 上記1に関連して、それぞれは意味のある最適化であるが、それらが同時に動作するためには、何らかの制約または制限を加えることをしないと、かえって性能が落ちる場合がある。

【0013】4. プログラムの解析は今までも行われていたが、その解析のレベルが低く、ループ最適化後のコードは最適な状態には至っていない。例えば、今までのアンローリング数の決定は、種々の最適化が実施される前のループ内部の変数の依存関係を解析するため、最適コードを生成するには限界があった。したがって、ループ内変数の依存関係から実施できる最適化を予想し、その上でその最適化が行われた後のループ内で使用される変数の依存関係を把握して使用レジスタ数を見積もることによりアンローリング数を決定する必要がある。

【0014】5. 上記4で示したプログラム解析の解析レベル制約が高度でない。プログラム解析によって得られる情報としては、最適化実施後の変数の依存関係を意味した使用レジスタ数、実行時間重みの付いた演算数等様々な情報がある。これらの情報を翻訳レベルを参照して獲得するかしないかを決定することによって、プログラムの解析レベルを段階的に制御することができる必要

がある。

【0015】本発明は上記問題点の解決を図り、複数のループ最適化を個々に実施するのではなく、プログラム／ハードウェアの特性をオブティマイズレベルに応じて解析し、実施する値のあるループ最適化をより性能が上がる条件で、かつより性能の上がる組合せで実施できるようにすることを目的とする。

【0016】

【課題を解決するための手段】図1は本発明の原理説明図である。図1において、10はCPUおよびメモリ等からなる処理装置、11はソースプログラムをオブジェクトプログラムに翻訳するコンパイラ、12はソースプログラムの入力および初期処理を行うフロントエンド部、13はプログラム構造を解析するプログラム構造解析部、14は最適化処理を行うループ最適化部、15は変数等に対してレジスタを割り付けるレジスタ割り付け部、16は命令実行順序等のスケジューリングを行う命令スケジューリング部、17はスケジューリング結果に従ってオブジェクトコードを生成するコード生成部を表す。また、20はループ最適化制御部であって、オプション情報解析部21、ハードウェア情報解析部22、プログラム解析部23、ループ最適化制御決定部24からなるもの、30はループ変形部であって、ループ・アンローリング、ループ・リダクション（詳しくは後述する）、ソフトウェア・パイプラインニング等の最適化処理を行うものを表す。

【0017】フロントエンド部12、プログラム構造解析部13、レジスタ割り付け部15、命令スケジューリング部16およびコード生成部17の処理については従来と同様でよいので、詳しい説明は省略する。

【0018】本発明では、ループ最適化部14のプロローグ部にループ最適化制御部20を設けることによって、ループ内で実施する値のある最適化を選択／制御しておき、ループ変形部30にその旨を伝達することにより、任意のプログラム／アーキテクチャ上で、翻訳／実行性能を向上できるループ最適化のみを実施する。

【0019】従来の最適化では、最適化同士がお互いに干渉し合うことがなく、2つの最適化が同時に動作することにより性能が低下することがあったが、この発明により効果的に最適化を動作させることができる。

【0020】図1に示すように、ループ最適化制御部20は、オプション情報解析部21、ハードウェア情報解析部22、プログラム解析部23、ループ最適化制御決定部24に細分化される。

【0021】オプション情報解析部21／ハードウェア情報解析部22では、指定されたハードウェア／オプション環境をプログラム解析部23／ループ最適化制御決定部24に伝達する。例えば、オプション情報として、オブティマイズレベル、各最適化の実施／抑止等の情報、また、ハードウェア情報として、レジスタ数、キャ

ッシュサイズ、ターゲットのマシンが持つ各命令の実行時間等の情報がある。これらの情報を、あるテーブル上に記録し、後の処理過程にメッセージとして伝達する方法をとる。

【0022】プログラム解析部23では、オプション解析により決定したオブティマイズレベルに依存して、どの程度のプログラム解析をするのかを決定する。プログラム解析で収集する情報の例として以下の3つがある。

【0023】1. ループ内に出現する演算数を各演算の実行時間重みを加味してカウントした値。

2. ループ内で使用される変数間の依存関係を把握し、実際に実施されるときと同じ条件かあるいはそれより緩い条件で、ループ変形部30で実際に行われる最適化を予想した結果。

【0024】3. その予想した最適化が実際に実施されることを加味した上で、ループ内で使用されるレジスタ数をループに対してローカルに必要な整数型変数／浮動小数点型変数、ループに対してグローバルに必要な整数型変数／浮動小数点型変数に分けて見積もった値。

【0025】上記2において、変数の依存関係を把握する方法の一例としては、ループ内の演算を各ノードとして作成したダグによる解析法がある。上記3において、最適化予想を加味する理由は、ループ・リダクションその他のループ最適化により、ループに対してローカルな変数が増加するためである（リネーミング以外の最適化結果を加味している点が今までと違うところである）。

【0026】例えば、これらの1, 2, 3の情報の収集をオブティマイズレベルに依存して可／不可にすれば、プログラムの解析レベルを操作することが可能になる。さらに具体的には、オブティマイズレベルが低い場合には、作成できるロードモジュールの実行時間をそれほど追求する必要がないので、翻訳性能が上がるように1のみの収集を行う方法をとれば、実行性能もある程度維持でき、かつ、翻訳性能が上がる最適化を導くことができる。

【0027】ループ最適化制御決定部24では、オプション情報解析部21、ハードウェア情報解析部22、プログラム解析部23からの一連の情報をもとに実施可能な最適化を決定する。そして同時に、プログラム／アーキテクチャの特性を把握し、同時に動作させるべきでない最適化を優先度に従って抑止、あるいは制限付きで実行させるようにループ変形部30に伝達する。この伝達の方法の一例としては、どのような最適化をどのように実行させるかを記したテーブルを作成し、それをメッセージとしてループ変形部30に伝達する方法がある。

【0028】ループ最適化制御決定部24の情報の一つであるアンローリング数は、プログラム解析部23での解析レベルに依存して決定される。最適化予想とそれを加味したレジスタ数が算出されていれば、ループ変形部

7

30で行われる最適化を十分に活かし、より並列化効果のアンローリング数を算出できる。アンローリング数の決定方法の実施例は後述する。

【0029】ループ変形部30では、授受したその情報を参照しながら、ループ変形をする種々の最適化、例えば、ループ・アンローリング、本発明に係るループ・リダクション、リネーミング、ソフトウェア・パイプライン等を実施する。

【0030】続いて、ループ・リダクションの実現方法について述べる。ループ・リダクションは、ループ中で他の演算と依存関係のない変数Qが、 $Q = Q + \alpha$  ( $\alpha$ は定数でない)という形で出現し、かつ、アンローリングが行われた後に行われる特殊なリネーミング処理である。アンローリング後、ループ内に含まれる $Q = Q + \alpha$ という同じ演算 ( $\alpha$ が配列の場合には配列の添字が異なる)に対して、ある1つの演算中の変数Qを除いたすべての変数Qをリネーミングすることにより、Qを含む演算間のdependencyをなくすることができる。これにより、特に1サイクルで複数の命令を実行できるプロセッサでは並列化が大きく促進される。ただし、Qをリネーミングしてしまつたため、ループ出口における変数Qは総和演算結果とならない。したがって、ループの前にリネーミングした変数Q $\gamma$  ( $\gamma = 1, 2, 3 \dots$ )を初期化する処理、ループの後に変数Qに対して変数Q $\gamma$  ( $\gamma = 1, 2, 3 \dots$ )を加算する処理が必要となる。このループの後に出す変数Qに対する補正処理は、 $Q = Q + Q1, Q = Q + Q2, Q = Q + Q3, \dots$ というような演算ではなく、余分なレジスタを使用せずに並列化を促進するように、Tree Height Reduction (演算木の高さをできるだけ小さくし、並列化を促進する最適化)を施す必要がある。ただし、ループ・リダクションは総和演算の順序を変更するため、実行結果に誤差を生じる可能性のある最適化である。

【0031】

【作用】従来技術では、例えばA, B, C, ...という複数の最適化項目があったとき、まず最適化項目Aについての実施可否/実施条件の判定、実施済の場合に最適化項目Aの実施、次に最適化項目Bについての実施可否/実施条件の判定、実施済の場合に最適化項目Bの実施、次に最適化項目Cについての実施可否/実施条件の判定、実施済の場合に最適化項目Cの実施、...というようにそれぞれ個別に最適化を実施していたのに対し、本発明では、ループ最適化制御部20を設けることにより、まず最適化項目A, B, C, ...のすべての実施可否/実施条件を最初に総合的に判定し、最適化の実施予想から最も効果的な実施条件、実施順番等を決めて、ループ最適化を行う点が大きく異なる。

【0032】本発明により以下の事項が可能となる。

・ループ最適化を予想してから、ループ最適化の制御をすることができる (最適化実施後のループ内の依存関係

8

を把握できるため、高度なプログラム解析および高度な最適化制御ができる)。

【0033】・オブティマイズレベル等によりプログラム解析のレベルを制御できる (それに依存して、高いオブティマイズレベルに対してより並列化効果の高いアンローリング数を決定できる)。

【0034】・上記のように、アーキテクチャ/プログラムをコンパイラが高度に解析することにより、動作して性能の上がる最適化のみを実施することが可能になる。

・また、同時に動作するとかえって性能が低下するような最適化同士を優先度に従い抑止、あるいは制限付きで実行することが可能になる。

【0035】・ループ・リダクションにより並列化が大きく促進する。

これらの機能 (制御) によるより豊富な情報をもとに、種々の最適化をより効果が上がるように実施/抑止し、翻訳/実行性能を向上させることができる。

【0036】

【実施例】図2は、本発明に係るループ・リダクションの例の説明図である。本発明では、ループ最適化の一つとして、ループの並列度を高めるための新たなループ変形を採用する。これをループ・リダクションと呼ぶ。ループ・リダクションは、例えば、図2(A)に示すようなループ・アンローリングの結果に対して、ループ中で他の演算と依存関係のない変数Qを含む演算のうち、その2番目以降の演算における変数Qを、図2(B)に示すように、それぞれQ1, Q2, Q3とリネーミングし、Qを含む演算間の相互依存性をなくすることにより、Q, Q1, Q2, Q3の演算に関する並列実行性を促進する最適化である。

【0037】ループ・リダクションは、プログラムの構造を変えないで単に変数名だけを他の変数名に置き換える従来のリネーミングとは異なり、元々は同一に扱われるべき変数を異なる変数として扱うものであるから、ループの前に値の初期処理、ループの後に補正処理のような付加的な処理が必要になる。

【0038】図3は、本発明の実施例によるループ最適化部14の各フェーズを詳細化したフローチャートである。ループ最適化部14に入力されたプログラムは、ループ最適化制御部20に入る。そして、オプション情報解析部21でオブティマイズレベル、最適化の実施抑止オプションを解析して各オプションの情報を一覧できるオプション情報テーブルを作成する。一般的に、あるオプションがオプションに連動する (実際には最適化自身が連動するのであるが) ことはよくある。例えば、オブティマイズレベルを示すオプションが良い例で、オブティマイズレベルを高いレベルにすると、コンパイラはそのプログラムに対して有効と思われる最適化をできるだけ動かす。したがって、あるオプションによって連動

9

(抑止)される最適化のチェックが必要となる。すなわち、オプション解析では単に指定されたオプションの認識だけでなく、それから連動される最適化を解析する。最適化レベルが最低レベルでループ最適化を行う必要がない場合には、図3のオプション情報解析部21の下での最適化レベルに関する判定でループ最適化部14から脱出する。

【0039】ハードウェア情報解析部22では、ロードモジュールを実行するマシンの様々な情報、例えば、整数型/浮動小数点型のレジスタ数、キャッシュサイズ、各アセンブラ命令の実行時間などを、そのマシンのハードウェア情報としてハードウェア情報テーブルに登録する。採取する情報量は最適化レベル等に依存して変化する。

【0040】プログラム解析部23では、オプション情報解析部21で得られた情報(最適化レベル)に依存して、プログラム解析のレベルを変動する。プログラムで解析すべき情報としては、例えば以下のものがある。

【0041】1. ループ内に出現する演算に対して、ハードウェア情報テーブル上に記録された各演算の重みを加味した実行時間をカウントした値(実行時間重み付き演算数)。

【0042】2. ループ内で使用される変数間の依存関係を把握し、かつオプション情報テーブルを参照することによって実施許可が与えられていることを考慮して、ループ変形部30で実際に行われる最適化を予想した結果。

【0043】3. その予想した最適化が実際に実施されることを加味した上で、ループ内で使用されるレジスタ数をループに対してローカルに必要な整数型変数/浮動小数点型変数、ループに対してグローバルに必要な整数型変数/浮動小数点型変数に分けて見積もった値。ここで、最適化予想を加味する理由は、リネーミング、ループ・リダクションといったループ最適化により、ループに対してローカルに必要な変数が増加するためである。

【0044】これらのどの情報が得られるかを最適化レベルに依存して変化させることによって、プログラムの解析レベルを変動することができ。上記の例では、項目1に対して、項目2,3より高度な解析レベルで実施することになる。プログラム解析部23でもプログラム情報テーブルが作成され、上記の情報が記録\*

\*される。

【0045】ループ最適化制御決定部24では、オプション情報テーブル、ハードウェア情報テーブル、プログラム情報テーブル、およびこのループ最適化制御決定部24が持つ最適化の優先順位テーブルを参照して、どの最適化をどのような順序でどのように実施するかを決定する。この最適化優先順位テーブルには、各最適化に対して、その優先順位と、自分より優先順位の高い最適化が動作したときに抑止されるか否かのフラグ、また、その最適化が実施されるための条件(例えば、ループ・リダクションであったならばループ・アンローリングが実施されていること)等が記録されている。

【0046】ループ最適化制御決定部24では、プログラム情報テーブルから予想される最適化を認識し、ループ制御情報テーブル中に実施か抑止かを登録する。ただし、その設定の際に、最適化優先順位テーブルを参照し、同時に実行すべきでない最適化を抑止、あるいは制限付きで実行するようにする。例えば、ループ・アンローリングとループ・リダクションが予想された場合、最適化優先順位テーブルを参照して、ループ・アンローリングを制限付きで実施するという情報を得る。それからダグ情報から得たループ変形後のレジスタ数を参照して(ループ・リダクションによりループに対してローカルに必要な変数が増えていることが考慮されている)、以下の計算式(1)によりアンローリング数を算出する。

【0047】関数F UNCでは、ループ内で行われるループ最適化を加味して、第1引数、第2引数で与えられるアンローリング数のチューニング値を返す。例えば、ソフトウェア・パイプラインが予想されたループの場合には、ターゲットであるマシンの浮動小数点型の四則演算の中(除算を除く)で最大の実行時間の倍数である値を返し(ただしUNROLL1は越えない)、ループ・リダクションが予想された場合には、下の式(2)でリダクションを加味して求められた値UNROLL1をそのまま返すというようなチューニングをする。また、ソフトウェア・パイプライン、ループ・リダクションが予想されるときアンローリング数はレジスタ数を越えず、さらに、それらが予想されない時のループのアンローリング数でも、アンローリング後のループ内の総演算数がキャッシュサイズを越してしまわないように考慮している。

【0048】

アンローリング数

=FUNC (UNROLL1, UNROLL2, OPT) (1)

UNROLL1

=MIN [ { (INUM-ILocal) / IGlobal },  
{ (RNUM-RLocal) / RGlobal } ] (2)

UNROLL2

= (MAX\_IST) / IST (3)

・FUNC 0 : 最適化OPTが行われることを加味し

て、第1引数、第2引数で与えられるアンローリング数



11

をチューニングしてその値を返す。

【0049】・UNROLL1:最適化後のループ内の使用変数を見積もって式(2)から決定したアンローリング数。UNROLL1は、基本的にソフトウェア・パイプライン、ループ・リダクションの実施に対して、ループ内レジスタの再利用がないことを考慮して見積もった値である(これらの最適化が実施されない時は、アンローリング後、レジスタの再利用が行われることが多い)。

【0050】・UNROLL2:ループ内の各演算に対して、それぞれの実行時間を重みとして加味して算出した総演算数から式(3)により決定したアンローリング数。

・MAX\_IST:アンローリング後のループ内の時間重み付き総演算数の最大値。

【0051】・IST:ループに対して算出した時間重み付き総演算数。

・OPT:ループ変形をするアンローリング以外のループ最適化(ex.0x0002)。

【0052】・INUM:システムが持つ整数型レジスタ数。

・RNUM:システムが持つ浮動小数点型レジスタ数。

・LOCAL:ループに対してローカルに必要な整数型レジスタ数。

【0053】・IGLOBAL:ループに対してグローバルに必要な整数型レジスタ数。

・LOCAL:ループに対してローカルに必要な浮動小数点型レジスタ数。

・GLOBAL:ループに対してグローバルに必要な浮動小数点型レジスタ数。

【0054】以上の関数FUNCによって得られるアンローリング数もループ制御情報としてループ制御情報テーブルに登録される。ループ変形部30では、そのループ制御情報テーブルを受け取り、それらの情報に従って種々の最適化を実施する。例えば、ループ・アンローリング/実施/5回展開、ループ・リダクション/実施、ソフトウェア・パイプライン/抑止というようなテーブル上の情報により、種々の最適化を実施/抑止する。

【0055】図4は本発明の実施例による最適化の例を示す図、図5は本発明の実施例によるループ最適化部の動作例の説明図、図6および図7は本発明の実施例において用いられるテーブルの例を示す図である。

【0056】以下に、本発明の実施例によるループ最適化部14が実際のプログラムの入力に対してどのように動作するかを、図4(A)に示すプログラムA(Progra A)を例にとって説明する。

【0057】コンパイラに対してプログラムAを与えらるると、プログラムAは図1に示すフロントエンド部12、プログラム構造解析部13を通過した後、ループ最

12

適化部14に入力され、さらに、ループ最適化制御部20に到達する。

【0058】図5に示すように、ループ最適化制御部20では、まず、オプション情報解析部21が動作し、ここでは、オプション指定により高いオプティマイズレベルを要求されていることを認識し、また、リネーミング抑止オプションが指定されていることを認識する。そしてさらに、高いオプティマイズレベルで運動する最適化を認識する。これらの情報を後のフェーズに伝達するために、図6(A)に示すようなオプション情報テーブルを作成する。

【0059】オプション情報テーブルにおけるオプティマイズレベルは、次のように分別される。

0:最適化しない

1:低レベルの最適化

2:中レベルの最適化

3:高レベルの最適化

また、オプション情報テーブルにおける最適化は、次のように分別される。

【0060】

0x0001:ループ・アンローリング

0x0002:ループ・リダクション

0x0004:ソフトウェア・パイプライン

0x0008:リネーミング

すなわち、0x0007は、ループ・アンローリングとループ・リダクションとソフトウェア・パイプラインが対象になることを示す。

【0061】ハードウェア情報解析部22では、オプション情報解析部21の情報(オプティマイズレベル)をもとに収集する情報を決定し、ハードウェア情報テーブルに登録する。いま、ターゲットとなるマシンの整数型のレジスタ数が32、浮動小数点型のレジスタ数が32、キャッシュサイズ16Kbyteとする。これに対して、図6(B)に示すようなハードウェア情報テーブルが作成される。

【0062】プログラム解析部23では、図6(A)に示すオプション情報テーブルを参照し、高いオプティマイズレベルが要求されていることを認識し、最も高度なプログラム解析を行う。図5に示すように、まず、ループ内の各演算に対して各演算の実行時間をその演算の重みとして付加したループ内の総演算数を算出する。続いて、ループ内の各演算の依存関係を解析するために、ループに対して、各演算をノードとするダグを作成する。このダグから、プログラムAのループがループ・リダクションできる依存関係であることを認識する。また、ループ・アンローリングできることも認識する(ループの出口が一つであること等を確認する)。そして、オプション情報テーブルを参照して、その最適化が抑止されないかを確認し最終的な予想結果を得る。予想される最適化は、実際に実施される最適化をすべて含む。した

13

が、予想されない最適化が動作することはない。

【0063】次に、それらの最適化を加味した上で、ループ内で使用されるレジスタ数を見積もる。ループ内で使用されるレジスタ数は、ループ内ローカルな整数型/浮動小数点型、ループ内グローバルな整数型/浮動小数点型に分けて見積もる。これらの情報は、図6(C)に示すようなプログラム情報テーブルに登録される。

【0064】ループ最適化制御決定部24では、ハードウェア情報テーブル、プログラム情報テーブルを参照して、ループ・アンローリング、ループ・リダクションが10 予想されていること等を認識する。そして、図7(A)に示すような最適化優先順位テーブルを参照し、個々の最適化がどのように実施されるかを決定する。その結果を、図7(B)に示すようなループ制御情報テーブルに記録する。最適化優先順位テーブル中の実施条件は、テーブルの第1列に並ぶ最適化が動作する前に必ず動作していなければならない最適化を示す。

【0065】プログラムAでは、ループ・アンローリングとループ・リダクションが予想されているので、実施順序はループ・アンローリング、ループ・リダクション20 実施条件はループ・リダクションを実施できるためのループ・アンローリングが実施されているから合格となる。また、ループ・アンローリングはループ・リダクションの実施のため、ループ・アンローリングが単独で動作したときよりもアンローリング数が少なく抑えられことになる。

【0066】最適化優先順位テーブル中の制限付き実施について説明する。

(ループ・アンローリング) テーブル上にあるような並列化効果のある最適化が予想されている時のループ・アンローリング数は、ループ・アンローリングが単独で実施された場合と比較して、リネーミング処理によりレジスタ数が増えているため、その数は少なく抑えられてい30 ます。

【0067】(ループ・リダクション/ソフトウェア・バイプライニング) ループ・リダクション、またはソフトウェア・バイプライニングの実施とループ・アンローリングの実施が同時に予想されても、アンローリング数がある閾値(それぞれの最適化により異なる)より小さく、並列化効果が期待できないと判断された場合には、アンローリング以外の最適化を抑制することがある。

【0068】また、ループ・リダクション、ソフトウェア・バイプライニングが同時に予想される場合には、アンローリングも当然予想されている、基本的には、ル\*  
式(1)の数値=UNROLL1-UNROLL1%BASE (4)

・BASE: ターゲットとなるマシンにおける浮動小数点型の除算を除く四則演算(ex. fadd, fmul, fsub)の実行時間の最大値。

【0072】A%B: AをBで割った余り。また、各最適化の実施順序は、最適化優先順位テーブルとプログ50

14

\*ループ・リダクションを優先し、ソフトウェア・バイプライニングを抑制するが、上記の理由で、ループ・リダクションが抑制された場合でソフトウェア・バイプライニング可能な場合にはソフトウェア・バイプライニングのみ実施する。

【0069】アンローリング数は、プログラム情報テーブル上の各レジスタ数を先述した式(1)の計算式に代入することにより求められる。これにより、OPTで与えられる最適化を十分意識した並列効果を最大にできるアンローリング数(6回)が算出される。このアンローリング数は、図7(B)に示すように、ループ制御情報テーブルのループ・アンローリングの備考欄に記録される。プログラムAでは、式(1)の関数は、OPTで与えられるループ・リダクションの実施予想を識別して、すでにループ・リダクションによりループに対してローカルに必要な変数が増えることを考慮して見積もったレジスタ数から算出済みであるアンローリング数UNROLL1をそのまま返す。ただし、このUNROLL1がある閾値より小さかった場合には、ループ制御情報テーブルのループ・リダクションの欄を抑制にし、ループ・リダクションが予想されなかったとして、アンローリング回数のチューニングを続け最終的に求まった値を関数値として返す。

【0070】式(1)のOPTに他の最適化がきた場合のアンローリング数のチューニングの実施例を挙げておく。ソフトウェア・バイプライニングが予想された場合のアンローリング数は、浮動小数点型の除算を除く四則演算の実行時間の最大値の倍数、かつUNROLL1を越えない値となる(下記の式(4)参照)。ただし、リダクションの場合と同様、アンローリング数UNROLL1がBASEより小さい場合にはループ制御情報テーブルのソフトウェア・バイプライニングの欄を抑制にし、ソフトウェア・バイプライニングが予想されなかったとして、アンローリング回数のチューニングを続け最終的に求まった値を関数値として返す。また、リネーミング効果がある場合にはリダクションと同様であるが閾値はなくUNROLL1がそのまま関数値となる。それ以外30 は、式(1)のUNROLL1とUNROLL2に対してある閾値を設け、UNROLL1がその閾値を越えたら両者の最小値、UNROLL1とUNROLL2がその閾値以下であつたら両者の最大値、それ以外はUNROLL1を返す。

【0071】(ソフトウェア・バイプライニングが予想されたときのアンローリング数)

ラム情報テーブルを参照して記録される。ただし、算出したアンローリング数が小さい場合の最適化抑制の情報はすでにループ制御情報テーブルに記録されているのでその情報も参照する。この結果、プログラムAでは、ループ・アンローリング、ループ・リダクションという実

15

行順序になる（ループ・アンローリングの制限は、上記のアンローリング数算出の際にすでに加味してあるので、実施の順序さえ伝達すれば良い）。

【0073】ループ変形部30は、ループ制御情報テーブルを参照し、ループ・アンローリングとループ・リダクションの実施が許可され、その順序で最適化を実施することを認識する。また、アンローリング数6回が最適であるという情報をそのテーブルから得る。これにより、図4（B）のように、まず、6重展開のアンローリングが行われ、その後でループ・リダクションが行われる。これにより、ループ・リダクションの効果が最も期待できるアンローリング数でループを展開することができ、ループ内の演算の並列度も最大になる。続いて、ループ・リダクションを図4を例にとって説明する。ループ・リダクションは、図4中のプログラムAの $Q=Q+A(I)$ というパターンを含むループがアンローリングされたとき、その2つ目以降のQに対して $Q \rightarrow Q \cdot \alpha$ （ $\alpha$ はアンローリング数（6）-1）という特殊なリネーミングを行う最適化である。これにより、Qを含む演算と $Q \cdot \alpha$ を含む演算の依存関係がすべてなくなり並列化が促進される。ただし、この補正処理として、ループ前に $\Sigma Q \cdot \alpha = 0$ という演算を入れ、ループの後に $Q = Q + \Sigma Q \cdot \alpha$ を入れる必要がある（ $\Sigma$ は $\alpha = 1$ から5までの総和）。また、このループの後の補正演算は、Tree Height Reductionにより図4（C）に示すループ・リダクション後の命令列の後部における加算演算のように、演算の高さができるだけ低くなるような形に変形される。

【0074】前述したアンローリング数のチューニング値を返す関数FUNCTIONのアルゴリズムの具体例を図8に示す。図8では、ループ・アンローリングが予想されていることが前提となっており処理を示している。図8において、SPはソフトウェア・パイプラインニング、retはFUNCTIONの関数値（算出されたアンローリング数）を表す。また、MAX（A, B）はA, Bの最大値、MIN（A, B）はA, Bの最小値を表す関数である。

【0075】①まず、リダクションが予想されたループに対する処理として、リダクション可で、かつUNROLL 1以上であれば、UNROLL 1を関数値として返す。

【0076】②次に、ソフトウェア・パイプラインニングが予想されたループに対する処理として、ソフトウェア・パイプラインニングが可で、かつUNROLL 1が前述したBASE以上であれば、（UNROLL 1 - UNROLL 1 % BASE）を関数値として返す。

【0077】③リネーミングが予想されたループに対する処理として、リネーミング可であれば、UNROLL 1を関数値として返す。

④上記以外の場合、ノーマル・ループに対する処理であり、以下の処理を行う。

16

【0078】UNROLL 1が閾値2より大きければ、UNROLL 1とUNROLL 2のうち小さいほうを関数値として返す。

UNROLL 2が閾値2以下の場合、UNROLL 1とUNROLL 2のうち大きいほうを関数値として返す。

【0079】それ以外の場合にはUNROLL 1を関数値として返す。

以上のような処理によって、次のような作用効果がある。例えば従来のループ最適化にさらに本発明に係るループ・リダクションを加えた場合を想定する。

【0080】1. ループ内部の演算および使用変数が多く、かつ、ループ・アンローリングができる場合で、ループ・リダクションできるパターンがループ中に出現したとき、ループ内で使用される変数をリネーミングして、別の変数にしてしまうと、ループ内で使用される変数がさらに増加して、レジスタ割り当て処理でスビルを生成する可能性がある。本発明では、ループ最適化を予想して、事前に使用レジスタ数を見積もることが可能であるため、スビルの生成などを回避することができる。

【0081】2. 上記1に関連して、それぞれは意味のある最適化であるが、それらが同時に動作することによってかえって性能が落ちるような場合、すなわち、上記1の例で挙げたようなループに対して、ループ・アンローリング、ループ・リダクションを同時に実施すると、スビルを生成する可能性があるため、ループ・リダクションを抑止し、ループ・アンローリングのみを実行させるといったような何らかの制御が必要となる場合に、その適切な制御が可能になる。また、例えば、ループ・リダクションとソフトウェア・パイプラインニングが同時に実行予想されるとき、同時に動作するとかえって性能が落ちる可能性があるため、より並列化効果の高いループ・リダクションを優先させるような制御も可能になる。

【0082】3. 例えば、ループ・アンローリング、ループ・リダクションが実施されるループにおいて、同時に動作させる場合にアンローリングの数を制限するというような制御が可能になる。

【0083】4. 今までのアンローリング数の決定は、種々の最適化が実施される前のループ内部の変数の依存関係を解析するため、最適コードを生成するのには限界があった。これに対し、ループ内変数の依存関係から実施できる最適化を予想し、その上でその最適化が行われた後のループ内で使用される変数の依存関係を把握して使用レジスタ数を見積もることによりアンローリング数を決定することが可能になる。

【0084】5. 例えば、高いオブティマイズレベルを要求された場合は、解釈時間よりロードモジュールの実行時間を優先するためにできるだけ多くの情報を獲得しておき、低いオブティマイズレベルを要求された場合

は、翻訳時間を優先させるため、取得に時間のかからぬ情報のみを獲得することが可能になる。

【0085】上記項目の実現により、任意のプログラム、任意のハードウェアに対して、どの最適化をどのように実施したら並列効果の高いコードを生成できるのかをコンパイラ自身が把握できるようになる。

【0086】また、本発明の一つであるループ・リダクションは、ループ中に他の演算と依存関係のない変数Qが、 $Q = Q + \alpha$  ( $\alpha$ は定数でない)という形で出現し、かつアンローリングが行われた場合に施されるループ変形法であり、その並列化効果は非常に高い。

【0087】

【発明の効果】以上説明したように、従来は最適化同士が互いに干渉していなかったため、複数の最適化が同時に動作するとかえって逆効果になる場合があったが、本発明によれば、各最適化がループ内構造を解析することによって予想され、その最適化の実施を考慮に入れたループ内レジスタ数の見積りができ、また、実施予想された最適化が互いに干渉しあうことによって、最適化の効果が最大になるように種々の最適化を実施／抑止することができるようになり、ループに対してより最適なコードが生成できるようになる。さらに、オプション情報解析部、ハードウェア情報解析部、プログラム解析部、ループ最適化制御決定部の連携を良くすることによって、任意のマシン、任意のプログラムでより最適なコードを生成できるようになる。

【0088】ループ・リダクションの効果自体も大きく、ループ・リダクションが行われるときのループ内の演算の並列度が高くなるため実行性能は大きく向上する。

【図面の簡単な説明】

【図1】本発明の原理説明図である。

【図2】本発明に係るループ・リダクションの例の説明図である。

【図3】本発明の実施例によるループ最適化部の処理フ

ローチャートである。

【図4】本発明の実施例による最適化の例を示す図である。

【図5】本発明の実施例によるループ最適化部の動作例の説明図である。

【図6】本発明の実施例において用いられるテーブルの例を示す図である。

【図7】本発明の実施例において用いられるテーブルの例を示す図である。

【図8】本発明の実施例におけるアンローリング数のチューニング値を返す関数FUNCの処理説明図である。

【図9】従来のループ・アンローリングによる最適化の例の説明図である。

【図10】従来のリネーミングによる最適化の例の説明図である。

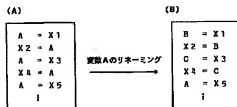
【図11】従来のソフトウェア・パイプライニングによる最適化の例の説明図である。

【図12】従来のソフトウェア・パイプライニングによる最適化の例の説明図である。

【符号の説明】

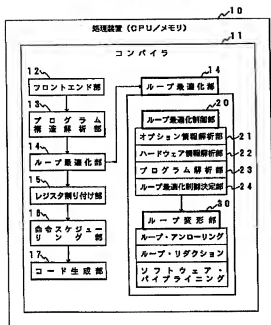
- 10 処理装置
- 11 コンパイラ
- 12 フロントエンド部
- 13 プログラム構造解析部
- 14 ループ最適化部
- 15 レジスタ割り付け部
- 16 命令スケジューリング部
- 17 コード生成部
- 20 ループ最適化制御部
- 30 オプション情報解析部
- 22 ハードウェア情報解析部
- 23 プログラム解析部
- 24 ループ最適化制御決定部
- 30 ループ変形部

【図10】



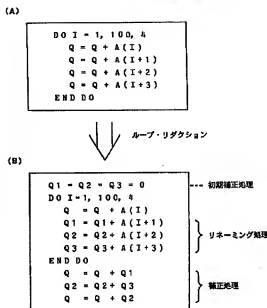
【図1】

本発明の原理説明図



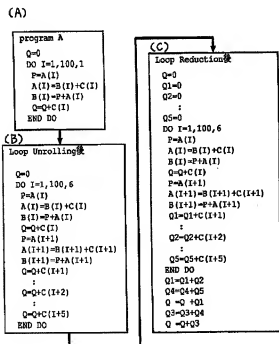
【図2】

ループ・リダクションの例の説明図



【図4】

最適化の例



【図6】

テーブルの例

(A)

オプションレベル	3
最適化	0x0007

(B)

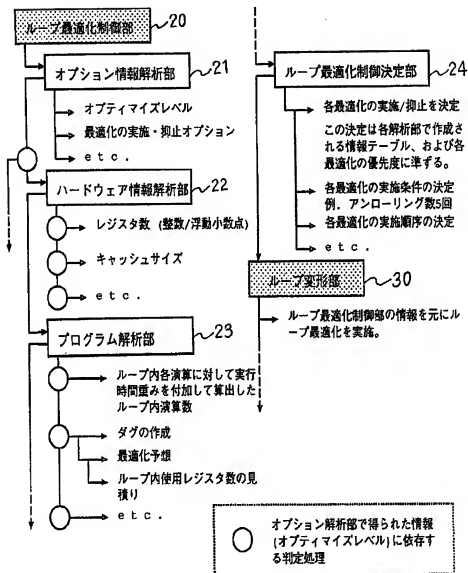
整数型レジスタ数	32
浮動小数点型レジスタ数	32
キャッシュサイズ	16Kbyte

(C)

手動最適化	0x0003
ループに対してローカルな変数型のレジスタ数	3
ループに対してグローバルな変数型のレジスタ数	4
ループに対してローカルな浮動小数点型のレジスタ数	0
ループに対してグローバルな浮動小数点型のレジスタ数	0

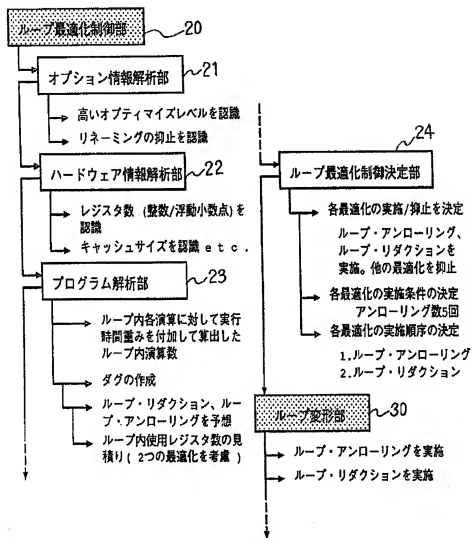
【図3】

## ループ最適化部の処理フローチャート



【図5】

## ループ最適化部の動作例



【図7】

テーブルの例

(A)

実施 順序	実施 条件	アロー リング	リダク ション	リネー ミング	バイプア イング
1	0=0000		制限付実施	制限付実施	制限付実施
2	0=0001	制限付実施		実施	実施
3	0=0000	実施	制限付実施	実施	実施
4	0=0001	制限付実施	制限付実施	実施	実施

(B)

最適化 ループ	実施順序	備考
ループ・アローリング	1	0(同)
ループ・リダクション	2	
リネーミング	0(停止)	
ソウトウェア・バイプライニング	0(停止)	

【図8】

FUNC の処理説明図

```

/* リダクションが予想されたループに対する処理 */
if (リダクション可 && UNROLL1 >= 閾値1)
    ret = UNROLL1;
else
    /* SP が予想されたループに対する処理 */
    if (SP 可 && UNROLL1 >= BASE)
        ret = UNROLL1-UNROLL1BASE;
    else
        /* リネーミングが予想されたループに対する処理 */
        if (リネーミング可)
            ret = UNROLL1;
        else
            /* 最適化が予想されないループに対する処理 */
            if (UNROLL1 > 閾値2)
                ret = MIN(UNROLL1, UNROLL2);
            else
                if (UNROLL2 <= 閾値2)
                    ret = MAX(UNROLL1, UNROLL2);
                else
                    ret = UNROLL1;

/* FUNC の戻数値 (アローリング版) の返却 */
return ret;

```

【図11】

【図9】

(A)

```

DO I = 1, 100, 1
  Q = Q + A(I)
END DO

```



ループ・アローリング

(B)

```

DO I = 1, 100, 4
  Q = Q + A(I)
  Q = Q + A(I+1)
  Q = Q + A(I+2)
  Q = Q + A(I+3)
END DO

```

(A)

```

11 LOAD
12 MULT
13 LOAD
14 MULT
15 ADD
16 LOAD
17 MULT
18 ADD
19 ST
21 LOAD
22 MULT
23 LOAD
24 MULT
25 ADD
26 ADD
27 MULT
28 ADD
29 ST
31 LOAD
32 MULT
33 LOAD
34 MULT
35 ADD
36 LOAD
37 MULT
38 ADD
39 ST
41 LOAD
42 MULT
43 LOAD
44 MULT
45 ADD
46 LOAD
47 MULT
48 ADD
49 ST

```

(B)

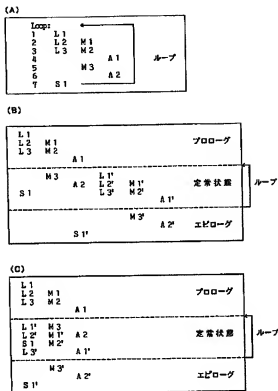
```

11 LOAD → L1
12 LOAD → L1
31 LOAD → L1
41 LOAD → L1
12 MULT → N1
22 MULT → N1
32 MULT → N1
42 MULT → N1
13 LOAD → L2
23 LOAD → L2
33 LOAD → L2
43 LOAD → L2
14 MULT → N2
24 MULT → N2
34 MULT → N2
44 MULT → N2
15 ADD → A1
25 ADD → A1
35 ADD → A1
45 ADD → A1
16 LOAD → L3
26 LOAD → L3
36 LOAD → L3
46 LOAD → L3
17 MULT → N3
27 MULT → N3
37 MULT → N3
47 MULT → N3
18 ADD → A2
28 ADD → A2
38 ADD → A2
48 ADD → A2
19 ST → S1
29 ST → S1
39 ST → S1
49 ST → S1

```



【図12】




---

フロントページの続き

(72)発明者 駒形 真  
 神奈川県川崎市中原区上小田中1015番地  
 富士通株式会社内